

LEKCE06

SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

vzorový výsledek cvičení

CVIČENÍ 6.1:

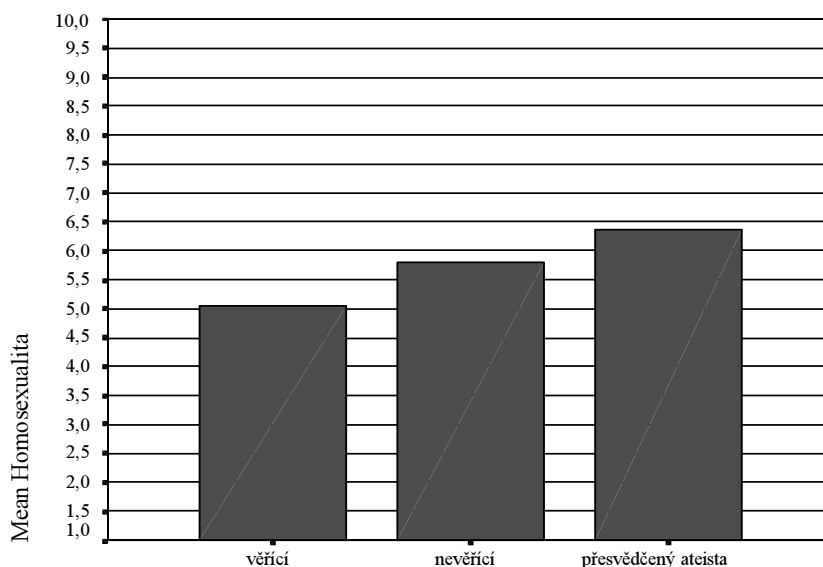
a) Zjistěte, zdali se podle dat EVS_ČR1999.sav liší postoj k homosexualitě (*q65a_8*) v závislosti na intenzitě náboženského přesvědčení (*q28*). Výsledek ilustrujte i graficky (upravte přitom měřítko osy Y tak, aby zachycovalo plnou škálu od 1 do 10). V odpovědi se zamyslete nad možnými důvody, které k tomuto výsledku vedly.

Považujme škálu za kardinální. V proceduře MEANS si vypočteme průměrné hodnoty dosažené na škále Q28 pro skupiny věřících, nevěřících a ateistů:

Report

Q65A_8 Homosexualita

Q28 Přesvědčení	Mean	N	Std. Deviation
1 věřící	5.05	739	3.58
2 nevěřící	5.79	840	3.49
3 přesvědčený ateista	6.36	141	3.45
Total	5.52	1721	3.55



Přesvědčení respondenta

Jak je zřejmé, postoje jednotlivých skupin se v našem souboru liší. Abychom mohli totéž tvrdit i o základním souboru, z něhož byl náš soubor vybrán, museli bychom testovat statistickou významnost vypočítaných rozdílů mezi průměry. V tomto případě by šlo buď o ONE-WAY analýzu nebo o tři t-testy, testující významnost rozdílů (víte ale, že mnohonásobné t-testy se nedoporučují a proto byste zvolili ANOVU –ONE WAY:

- věřící a nevěřící
- věřící a ateisté
- nevěřící a ateisté.

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

b) Spočítejte průměry tolerance k homosexualitě (*q65a_8*) pro skupiny respondentů podle jejich volebních preferencí. Výsledek zobrazte graficky a poté zjistěte, zdali je rozdíl v postoji k homosexualitě mezi stoupenci komunistů a stoupenci KDU statisticky významný. Použijte parametrického i neparametrického testu. V odpovědích vždy uvádějte všechna potřebná čísla jako argumenty. **POZOR:** Abyste byli schopni tuto úlohu vyřešit, budete muset z proměnné *q72* vytvořit pomocí RECODE novou proměnnou. Nazvěte ji *preferen* a její nové kódy vytvořte tak, aby měla tyto varianty:

1. KSČM,
2. ČSSD
3. KDU
4. US
5. ODS

Vše ostatní, včetně 90 nevolil by – missing value

Soubor s novou proměnnou uložte!

volební preference 99

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	KSČM	160	8,4	14,8	14,8
	ČSSD	286	15,0	26,4	41,2
	KDU	98	5,1	9,0	50,2
	US	163	8,6	15,1	65,3
	ODS	376	19,7	34,7	100,0
	Total	1084	56,8	100,0	
Missing	System	824	43,2		
Total		1908	100,0		

a) Spočítejte průměry tolerance k homosexualitě (*q65a_8*) pro skupiny respondentů podle jejich volebních preferencí. Výsledek zobrazte graficky a poté zjistěte, zdali je rozdíl v postoji k homosexualitě mezi stoupenci komunistů a stoupenci KDU statisticky významný. Použijte parametrického i neparametrického testu. V odpovědích vždy uvádějte všechna potřebná čísla jako argumenty.

Report

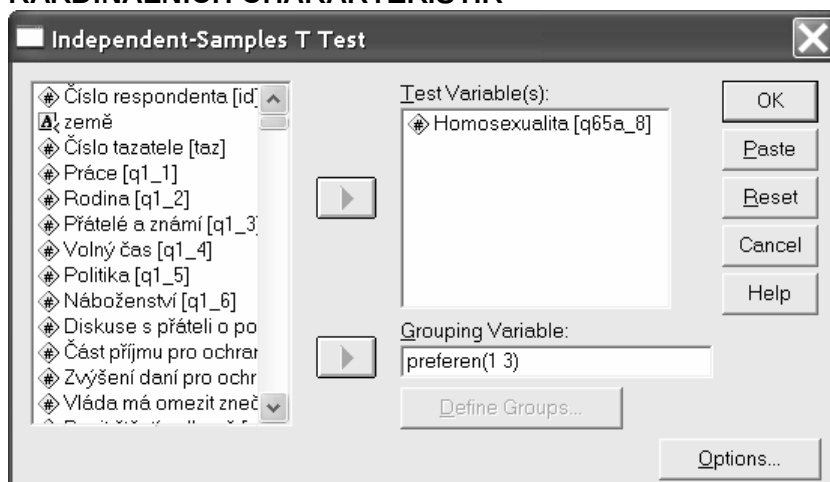
Homosexualita

volební preference 99	Mean	N	Std. Deviation	Std. Error of Mean
KSČM	3,95	156	3,310	,265
ČSSD	5,40	283	3,442	,204
KDU	4,17	97	3,521	,357
US	6,26	154	3,551	,286
ODS	5,72	356	3,635	,193
Total	5,31	1046	3,589	,111

Voliči KSČ jsou k homosexualitě nejméně tolerantní (průměr na škále je 3,95), ale voliči KDU-ČSL jsou jim, co se týče postoje k homosexualitě velmi podobní (průměr na škále je 4,17). Připomeňme si jen, že 1 znamená maximální netoleranci a 10 maximální toleranci v rámci dané škály. Voliči obou stran jsou v tomto ohledu méně tolerantní než voliči ČSSD, ODS a zejména Unie svobody (předbíháme, ale nemůže to být dáno tím, že je mezi nimi relativně nižší podíl osob s vysokoškolským vzděláním?).

Na místě je ovšem otázka, zda tento poměrně malý rozdíl nejde na vrub výběrové chyby, zda jsou rozdíly průměrné (ne)tolerance k homosexualitě skutečně statisticky významné. Použijeme tedy test statistické významnosti rozdílu (v tomto případě Independent Samples T-test)

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK



V Grouping Variable jsme definovaly skupiny, pro něž budou průměry porovnávány.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Homosexualita	Equal variances assumed	1,627	,203	-,489	251	,625	-,214	,439	-1,079	,650
	Equal variances not assumed			-,482	194,66	,630	-,214	,445	-1,092	,663

Podle výsledku F- testu volíme pro rozhodování první řádek (F - test nedovolil zamítnout nulovou hypotézu neboli hypotézu o shodě variancí - je zde 20,3% riziko chyby → Equal variance assumed). Také v t- testu jsme na tom stejně (riziko při zamítnutí nulové hypotézy o shodě průměrů je 62,5%).

Nelze tedy vyloučit, že rozdíl v průměrném umístění voličů obou stran mohl být dán výběrovou chybou a v základním souboru neexistuje. Zatímco v případě voličů KDU-ČSL můžeme předpokládat, že základem odmítnutí tolerovat homosexualitu vychází z jejich náboženské víry, v případě voličů KSČM bychom mohli předpokládat vliv jejich vysokého věku (mohli bychom to dále testovat) a celkové netolerantnosti – sblízuje je poměrně nižší vzdělání).

V modulu **NONPARAMETRIC TESTS: INDEPENDENT SAMPLES T-TEST** použijeme "Mann-Whitney" test:

Test Statistics^a

	Homosexualita
Mann-Whitney U	7963,500
Wilcoxon W	13013,50
Z	-,238
Asymp. Sig. (2-tailed)	,812

a. Grouping Variable: volební preference 99

Výsledek je stejný, test tentokrát upozorňuje na 81,2% riziko chyby při zamítnutí nulové hypotézy.

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

Zkusme jen voliče KSČM a voliče ČSSD (mají k sobě blízko mnoha postojů a rozdíl v toleranci k homosexualitě je druhý nejnižší).

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Homosexualita	Equal variances assumed	1,684	,195	-4,261	437	,000	-1,444	,339	-2,110	-,778
	Equal variances not assumed			-4,310	329,114	,000	-1,444	,335	-2,103	-,785

Podle výsledku F- testu volíme pro rozhodování první řádek (F - test nedovolil zamítnout nulovou hypotézu neboli hypotézu o shodě variancí - je zde 19,5% riziko chyby → Equal variance assumed). V t- testu ovšem můžeme nulovou hypotézu zamítnout (riziko, že zamítneme nulovou hypotézu přestože platí je menší jak 0,1%).

c) Pokusme se zjistit jak je to se statickou významností průměrné (ne)tolerance k homosexualitě mezi všemi skupinami voličů?

K tomu použijeme ONE-WAY analýzy

ANOVA

Homosexualita

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	612,291	4	153,073	12,390	,000
Within Groups	12849,000	1040	12,355		
Total	13461,291	1044			

Tato tabulka nám říká, že minimálně mezi dvěma srovnávanými skupinami voličů statisticky významný rozdíl v jejich průměrných hodnotách (ne)tolerance k homosexualitě existuje. Ještě ale stále nevíme mezi kterými. To nám řekne až Post Hoc Multiple Comparisons.

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

Multiple Comparisons

Dependent Variable: Homosexualita

Bonferroni

(I) volební preference 99	(J) volební preference 99	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
KSCM	ČSSD	-1,444*	,351	,000	-2,43	-,46
	KDU	-,214	,455	1,000	-1,49	1,06
	US	-2,302*	,399	,000	-3,43	-1,18
	ODS	-1,765*	,338	,000	-2,72	-,81
ČSSD	KSČM	1,444*	,351	,000	,46	2,43
	KDU	1,229*	,413	,030	,07	2,39
	US	-,858	,352	,148	-1,85	,13
	ODS	-,321	,280	1,000	-1,11	,47
KDU	KSČM	,214	,455	1,000	-1,06	1,49
	ČSSD	-1,229*	,413	,030	-2,39	-,07
	US	-2,087*	,455	,000	-3,37	-,81
	ODS	-1,550*	,402	,001	-2,68	-,42
US	KSČM	2,302*	,399	,000	1,18	3,43
	ČSSD	,858	,352	,148	-,13	1,85
	KDU	2,087*	,455	,000	,81	3,37
	ODS	,537	,339	1,000	-,42	1,49
ODS	KSČM	1,765*	,338	,000	,81	2,72
	ČSSD	,321	,280	1,000	-,47	1,11
	KDU	1,550*	,402	,001	,42	2,68
	US	-,537	,339	1,000	-1,49	,42

*. The mean difference is significant at the .05 level.

Jak ukazuje níže uvedená tabulka, existují statisticky významné rozdíly:

- mezi voliči KDU a voliči ČSSD, US a ODS
- mezi voliči KSČM a voliči ČSSD, US a ODS.

Neexistují však mezi voliči KSČM a KDU – není vyloučeno, že mají stejnou (ne)toleranci vůči homosexualitě.

Neexistují však mezi voliči ČSSD, US a ODS – není vyloučeno, že mají stejnou (ne)toleranci vůči homosexualitě.

Tyto 2 skupiny se v populaci svou mírou k homosexualitě odlišují.

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

CVIČENÍ 6.2:

Testujte, zdali výběrový soubor EVS_ČR1999.sav pochází ze základního souboru populace ČR z hlediska věku. Jelikož výběrový soubor zahrnuje populaci ČR starší 18 let, musíte nejdříve vypočítat průměrný věk tohoto základního souboru. Ve Statistické ročence České republiky 2000 si najdete tabulku Obyvatelstvo 4.-2 (str. 106) a z údajů ve sloupci 1. 7. 1999 Celkem (naš výzkum totiž proběhl v květnu 1999) a z řádků pro věk 18-99 ručně vypočtete průměrný věk populace starší 18 let (věk 99+ chápejte jako hodnotu 99). Tuto hodnotu pak použijte v SPSS.

Zatím řeším jen modelově, protože se mi podařilo na poslední chvíli zjistit, že jsem ztratil vypočítaný údaj. Dal jsem si ho zjistit a tak jen pro informaci uvádím řešení s odhadnutým věkovým průměrem (berte ho cvičně jako by byl zjištěn!).

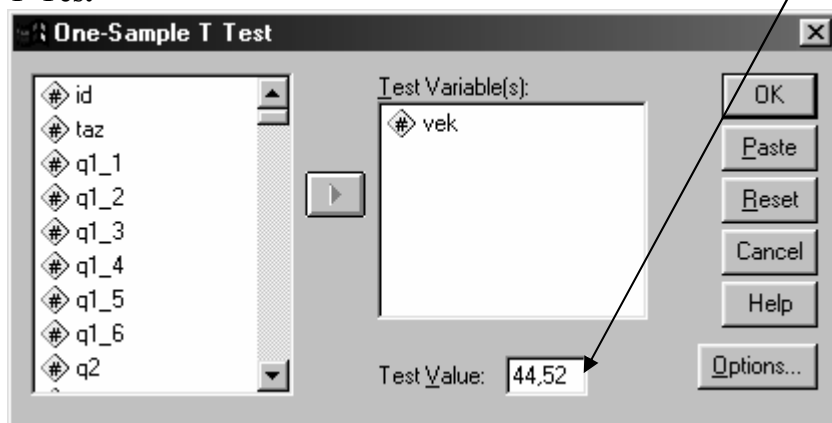
Průměrný věk v našem souboru je 45,69 let (průměrný věk mužů a žen v zadání nebyl požadován, ale logika SPSS si žádá použití alespoň jedné proměnné vymezující kategorie, pro kterou je průměr počítán (musí to být kategorizovaný znak, nejlépe o co nejmenším počtu kategorií).

Report

VEK			
Q84 Pohlaví	Mean	N	Std. Deviation
1 muž	44,53	916	17,101
2 žena	46,77	985	16,515
Total	45,69	1901	16,833

Předpokládejme, že jsme zjistili ve Statistické ročence průměrný věk 44,52 let:

T-Test



One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
VEK	1901	45,69	16,833	,386

One-Sample Test

	Test Value = 44.52					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VEK	3,034	1900	,002	1,17	,41	1,93

V tomto případě lze zamítnout nulovou hypotézu, věkový průměr v našem souboru se sice jen velmi málo, ale statisticky významně liší od věkového průměru základního souboru (riziko chyb při zamítnutí této hypotézy je jen 2%).

LEKCE 6: SROVNÁVÁNÍ SKUPIN NA ZÁKLADĚ STŘEDNÍCH HODNOT JEJICH KARDINÁLNÍCH CHARAKTERISTIK

CVIČENÍ 6.3

Odpovězte na otázky:

- a) Jaký je průměrný věk mužů a žen (Q84) v našem souboru a zda se tyto hodnoty od sebe liší i v populaci či je jejich rozdíl v našem souboru způsoben pravděpodobně jen výběrovou chybou.

Report

VEK

Q84 Pohlaví	Mean	N	Std. Deviation
1 muž	44.53	916	17.10
2 žena	46.77	985	16.51
Total	45.69	1901	16.83

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
VEK	Equal variances assumed	2.003	.157	-2.90	1899	.004	-2.23	.77	-3.75	-.72
	Equal variances not assumed			-2.89	1877.813	.004	-2.23	.77	-3.75	-.72

Průměrný věk je v našem souboru u mužů (44,53 let) a u žen (46,77). Tento rozdíl je statisticky významný, při zamítnutí nulové hypotézy o shodě těchto průměrů v populaci se dopouštíme jen 0,4% rizika chyby (čteme signifikanci t na prvním řádku).

- b) Jaký je průměrný věk mužů se základním vzděláním a žen se základním vzděláním v našem souboru a v jakém intervalu se pohybují tyto průměry v celé populaci?

Report

VEK

Q84 Pohlaví	VZDĚLÁNÍ kategorizace q94	Mean	N	Std. Error of Mean
1 muž	1 základní	38.77	174	1.30
	2 vyučen	45.65	418	.83
	3 SŠ	44.19	215	1.17
	4 VŠ	49.99	107	1.43
	Total	44.51	914	.57
2 žena	1 základní	51.67	192	1.44
	2 vyučen	47.66	362	.83
	3 SŠ	43.99	341	.81
	4 VŠ	43.19	89	1.42
	Total	46.77	985	.53
Total	1 základní	45.55	366	1.03
	2 vyučen	46.58	781	.59
	3 SŠ	44.06	556	.67
	4 VŠ	46.90	196	1.04
	Total	45.68	1899	.39

Při zadání výpočtu v proceduře MEANS přidáme proměnnou VZDĚLÁNÍ (Layer 2). Ženy se základním vzděláním jsou v našem souboru v průměru mladší než muži se základním vzděláním. Intervaly spolehlivosti nám v obou případech dovolí vypočítat zadaná standardní chyba průměru. Lze řešit i pomocí výpočtů v podsouborech mužů a žen (zejména chceme-li testovat statistickou významnost rozdílu - to ovšem nebylo zadáno).